Homework Assignment 2
601.467 / 667 Introduction to Human Language Technology
Fall 2019
Due: Saturday November 2, 2019

October 21, 2019

Language Modeling with RNNs

In this part of the assignment, you will implement character-level language model based on recurrent neural network.

Question 1. Open the following colab notebook and complete the assignment:

https://colab.research.google.com/drive/1lvBvTt$_u$K3DMLAKZU4XiRJ − 0ku6ydvHE

Follow all the steps specified in it. Include link to your solved notebook in your submission. Some parts of the notebook are optional and will not be graded.

Question 2. Explore different number of hidden states and layers in LSTM. Which one worked better?

Question 3. (Optional) Explore different different optimizers like SGD and ADAM and learning rates.

Question 4. Explain how the text generation process works.

1. How do you obtain probability distribution over the possible characters?
2. How do we incorporate a starting prefix during text generation?
3. How the temperature parameter affects the results? What will happen if the temperature parameter approaches zero? What will happen if the temperature parameter approaches Infinity? Paste examples from your colab notebooks.

Question 5. Explain why the sequential nature of recurrent neural networks makes them less efficient compared to methods based on convolutions, for example.

Question 6. In the homework exercise we have loaded whole text sequences into the GPU memory. Imagine that we will work with sequences of much greater length, how that will affect GPU memory consumption? How can you solve this problem? (One of the solutions is related to Question 9).

Question 7. (Optional) We have Implemented language model that can condition on the artist as well as a starting prefix.

1. Train and run this conditional RNNs with different artists (and some starting prefix). How does the text generation change with different artists? For example, use the prefix `hey hey` with artists `ABBA` and `Eminem`. Did the RNN capture any traits from the artist?

Question 9. (Optional) Implement truncated backpropagation through time. By what factor did it reduce the memory consumption?